

SOFTWARE AND HARDWARE SOUND ANALYSIS TOOLS FOR FIELD WORK.

Pavan G.^{1,2}, Manghi M.¹, Fossati C.¹

¹ Centro Interdisciplinare di Bioacustica e Ricerche Ambientali, Università di Pavia,
Via Taramelli 24, 27100 Pavia, Italy. Email gpavan@cibra.unipv.it

² Dept. of Urban Science, IUAV, Venice, Italy

1. ABSTRACT

The latest version of the real-time Digital Signal Processing Workstation developed at CIBRA runs in a standard Windows environment and can use a wide range of sound acquisition devices. It can be based on a notebook to allow on-field use. Depending on the acquisition devices, recording, analysis and display can be performed in real-time up to 500 ksamples/sec to provide useful bandwidth to more than 200 kHz. The software was primarily developed for continuous real-time monitoring in bioacoustical studies.

KEY WORDS: sound analysis, digital signal processing, spectrogram, real-time

2. INTRODUCTION

The development of DSP (Digital Signal Processing) techniques and of affordable high-speed computer hardware with large storage capacity has made the computer analysis and recording of bioacoustical signals an every-day tool for ethological research and for monitoring the underwater environment. Since digital techniques appeared and replaced analog systems in the recording and analysis of animal vocalizations, the advantages have not only been in the change of recording media but mostly in the development of completely new methodologies and strategies for analysis, interpretation and distribution of data. The exponential increase in speed and storage capabilities and the drop of costs were less predictable. The sum of these two factors makes it possible to have an entire bioacoustic laboratory in an affordable notebook, ready to be moved everywhere for recording and "seeing" infrasounds, sounds and ultrasounds.

Since 1980 the University of Pavia develops and applies digital techniques to the study of animal sounds. The latest version of the CIBRA's real-time Digital Signal Processing Workstation (DSPW) meets most of the needs for sound recording and analysis in the field as well as in the lab. Based on Windows, it delivers ease of use and full compatibility with standard sound devices and sound editing software. This makes easier the distribution of recordings among scientists and students.

2.1 HISTORICAL BACKGROUND

The development of a digital sound analysis system suited for bioacoustic research began at the University of Pavia in 1980. In those days the analysis of 1 second of sound required more than 40 minutes of computing, but the digital approach anyway proved to be a winning solution for its flexibility.

The first real-time portable system for field use was developed in 1990 by integrating a bulky portable PC (16MHz 386 CPU) and a DSP based signal acquisition board. The software was developed in DOS and it was mainly based on highly optimized assembler routines. It was used for

the first time in a cruise in 1991 and then presented at the XV IBAC meeting [1] and at the ECS meeting in Sanremo [2].

The continuous increase in CPU speed, in the following years, made the sound analysis faster and faster. The Intel 486/66 was the first CPU able to process in real-time up to 48000 samples/sec without expensive and hard to program DSP processors. Nowadays, a Pentium III 500 requires less than 100 msec to display a spectrogram of a 1 second signal sampled at 48000 samples/sec thus making it possible to theoretically work in real-time up to 480 K samples/sec.

Nevertheless, the CPU speed is one only of the factors affecting the performance of signal processing systems. Sound acquisition devices, bus speed, CPU speed, hard disk speed and other hardware and software related factors affect the sustained throughput of a system and its suitability for real-time application.

But what does real-time mean? In a general sense, a real-time analysis system is a system able to produce a result without appreciable delay. A more accurate definition, applied to acoustic systems, implies that the signal must be processed continuously, without gaps (without losing data), and without appreciable delay. In acoustic analysis, a real time system should show the features of an acoustic signal continuously without losing even short transients. Also, the display of features should happen simultaneously with the signal, or at least with a very short – but constant - delay.

3. SYSTEM DESCRIPTION

3.1 HARDWARE PLATFORM

The DSPW has been developed to provide acoustic researchers a compact, low cost, flexible and expandable personal system. The workstation is based on the Intel x86 platform and conforms to PC and Windows standards (earlier versions were developed in DOS [1-6]). It can be based either on desktop or notebook PCs. The software is highly optimized to exploit the processing capabilities of Intel Pentium processors and to perform smoothly in real-time.

3.2 SOFTWARE

The Windows software has been developed in 1997 to match CIBRA's research needs and it is continuously updated to meet research requirements and to take advantage of new hardware. In 1999 new functions were developed to meet the requirements of a research project granted by the Office of Naval Research.

The software package includes analysis, recording and display tools. Real-time spectrogram display, real-time cepstrogram display, wrap-around or scrolling display, wide control on all analysis parameters, frequency-time cursor while in real-time mode, frequency zoom capabilities, frequency tracking, recording and playback of files with real-time display, scheduled recording, on-event recording, file analysis are all implemented. Sound files and spectrograms can be saved in standard formats to allow further processing and managing with other software.

As the software has been developed primarily for analysis and display, editing capabilities are not implemented. These features can be easily found in a number of shareware and freeware.

3.3 SOUND ACQUISITION

To provide recording quality often better than DAT recorders many high quality sound acquisition boards are on the market. Several low-cost AD/DA boards are available for musical application and most of them can be used for sound acquisition. Nevertheless, not all of this equipment has enough acoustic quality or accuracy to be used in high quality recording and sound analysis.

The board should be accurately selected to match the needs of the research and the quality level of the whole electroacoustic chain used. Choosing boards with both analog and digital I/O is recommended.

Digital I/O is normally available in standards: electrical SPDIF on RCA connectors, electrical AES/EBU on XLR connectors and optical on TosLink connectors. Professional boards offer all three standards with the ability to convert one format to another, to work as a digital switching board, and to change SCMS protection codes.

As the software relies on the standard Windows multimedia interface, the software can work on almost any sound board on the market including digital I/O boards, 96 kHz sound boards, the still rare 192 kHz boards and the USB audio devices, unfortunately still limited to 48 kHz sampling. Almost all quality boards have 24 bit AD converters providing up to 110-115 dB of dynamic range, less than the theoretical limit of 144 dB because of the limitations of the analog front ends. Just a few external high-quality expensive converters can provide more than 120 dB of SNR.

The sound analysis software is currently limited to 16 bit signals, providing "only" 96 dB of dynamic range, well enough to work on a wide range of signals. Future version will overcome this limitation. In the meanwhile, it is possible to read 32 bit integer raw data files.

Multiple sound devices are supported to easily choose among the installed I/O options. Multiple program instances are allowed provided that each is using a different sound device (two programs can't access the same sound device as pipelining is not allowed in standard Windows drivers). Normally, high quality boards have both analog and digital I/O which can be used as independent devices. As the software allows independent selection of different devices for input and output, it is, for instance, possible to record from digital inputs and to play to analog outputs of the same board or of a different board. A great effort was put in testing a number of different devices and their combination to be sure of getting optimal performances on a wide range of platforms.

3.4 ULTRASONIC EXTENSION

Soundboards sampling at 96 kHz usually offer a 40-44 kHz bandwidth, while 192 kHz boards may provide up to 80-90 kHz bandwidth. If connected to suitable transducers, these boards may offer new chances to reveal and study ultrasounds. To cover the whole frequency range of both terrestrial and aquatic organisms, of those performing echolocation in particular, it is however required to extend the recording range up to 150 kHz at least. For this purpose more expensive, and often more noisy, sound acquisition systems are required. The ultrasonic extension of our software is tailored to use National Instruments DAQ cards, providing sampling frequencies of more than 1 MHz with 12 bit accuracy. This narrows the theoretical dynamic range to 72 dB, not a great range, anyway higher than the range of analog instrumentation tape recorders.

NI DAQ cards are available for PCI, PCMCIA (CardBus) and FireWire (IEEE 1394) computer bus. For desktop PCs, PCI cards are relatively cheap and perform well. For portable use on notebooks, PCMCIA CardBus boards are good because of their compact size but they are slower than the FireWire based acquisition systems, which are external devices larger in size and more expensive. On a PIII 500MHz notebook it is possible to acquire and analyze at 320 ksamples/sec (single channel) with a PCMCIA card and at about 500 ksamples/sec with a FireWire device. If recording to disk without displaying, these rates can be increased by 50%. (These rates cannot be guaranteed for every PIII 500 model).

3.5 RECORDING

Sound recording can be performed in two modes: RAM recording, which is limited by the available amount of RAM memory, and Hard Disk recording.

In RAM recording mode sounds are recorded in a buffer whose size is set by the user according to the available free memory. The buffer contents can be saved to disk when the acquisition is stopped or when the buffer is completely filled. A further mode allows to set the buffer as an endless circular buffer; when the user stops the acquisition the buffer holds the last *n* minutes of audio and the user can save them to disk.

To support long term monitoring and recording activities, the software has been improved to allow continuous operation for days. While displaying sounds, the software can continuously record to disk. To overcome the 2GB file size limit of Windows operating systems, the program creates a new file every hour (file length is anyway user selectable) and can use available disks in sequence. A buffering system allows recording continuously without loss of data; if something goes wrong and a data block is lost a warning and a log are generated.

Each file is stored with start date and time in the filename and a log of all recordings is generated automatically; if a GPS is connected to the DSPW, location data is added to filenames and log file.

Recording length is only limited by the available storage space. Storage space is normally limited by hardware constraints: the largest EIDE disks currently available hold 80GB each and standard desktop PCs allow no more than 4 hard disks. If sampling two channels at 48kHz, four 80GB disks can record continuously for 19-20 days.

By using two hard disks minimum, it is possible to backup a disk to high capacity SCSI tapes while the program is writing to the other disk(s). The backup procedure requires the user intervention.

If more storage space is required, it is possible to add additional controllers to increase the number of disks. For this purpose RAID EIDE controllers are inexpensive and very flexible, but limited to desktop PCs. More expensive SCSI controllers allow more than ten disks in line. In the near future even more flexible and powerful recording systems will be based on external hot-swappable FireWire disks to deliver huge storage capacity to notebooks as well.

Other than on the size, the choice of a disk should be based on the required throughput; for audio range there are no special requirements, but for ultrasonic recording, at rates of 1Mbytes/sec per channel, fast disks are required.

For very large amounts of data, in the order of hundreds of GB, there are no cheap and compact portable solutions and we look forward to the new optical technologies promising 10TB (maybe 100TB) of data on a single tick optical disk.

3.6 ANALYSIS

The spectrographic display is created calculating the FFT of signal frames extracted at regular intervals from the incoming signal. The FFT size is selectable ranging from 512 samples for the default display to 16384 samples for the zoom display.

Analysis parameters (sampling rate, FFT size, window length, window shape, scanning - or shift - step) can be individually modified and the software checks for parameter incompatibilities (if the user sets, for example, the window length greater than the FFT size, the FFT size is automatically increased to match the chosen window). It is also possible to set the window shorter than the FFT (zero padding) to increase time selectivity at the expense of frequency selectivity while maintaining a same frequency resolution. This makes the software very flexible, but requires a good understanding of sound analysis principles also in relation to the real processing capabilities of a given machine. The display structure has been designed to provide great accuracy and a precise relationship among displayed pixels and data resulting from FFT processing. A pixel in the spectrogram always reflects the value of a frequency bin of a given time window, unless averaging or zooming have been chosen by the user.

FFT sizes greater than 512 samples allow to select among packing frequency bins to fit the standard display or zooming into a frequency range (Figure 1) by directly plotting bins to the display. In this way the zooming factor corresponds to the chosen FFT size / 512 (a FFT size of 16384 yields a 32x zoom). When zoom is selected, it is possible to switch among zoom display and standard display while the spectrogram is running as well as to change the baseline of the zoomed range.

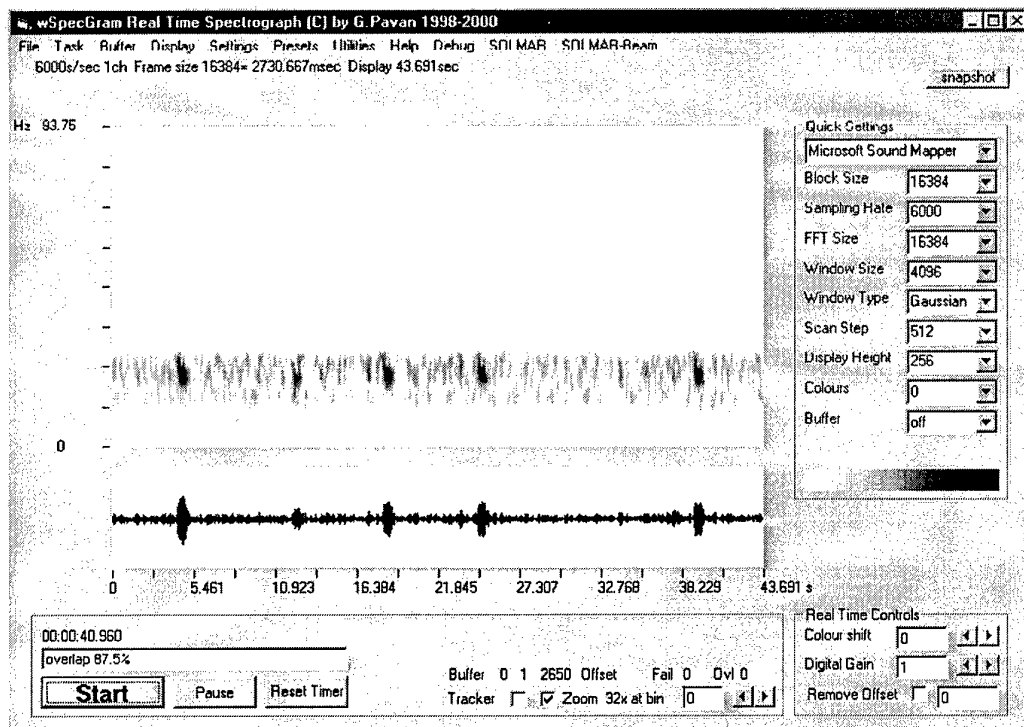


Figure 1 . Real-time spectrogram in 32x zoom mode revealing fin whale sounds at frequencies of about 20 Hz.

Other than the traditional spectrogram it is possible to display the cepstrogram, which is the plot of the cepstrum versus time. This display is particularly useful to analyze harmonic sounds and to improve the measure of their fundamental frequency. In marine mammal studies, the cepstrum is particularly useful to measure the Inter Pulse Interval (IPI) in sperm whale clicks [5].

Further real-time options are in development to provide filtering, band shifting and band compression to make easier the analysis and management of ultrasonic signals. Among these procedures, which are based on FFT manipulation and inverse transformation, band compression is the most critical because of the processing of phase information.

Figures 2-4 show that band shifting and downsampling can transform band limited ultrasonic signals to fit the frequency range of available recording instrumentation, to reduce data size and to make them audible while maintaining the original frequency-time features. The process has been made in separate steps for demonstration purposes, but it can be performed in a single phase in real-time. These procedures are temporarily implemented in separate routines and will be included in the main software soon.

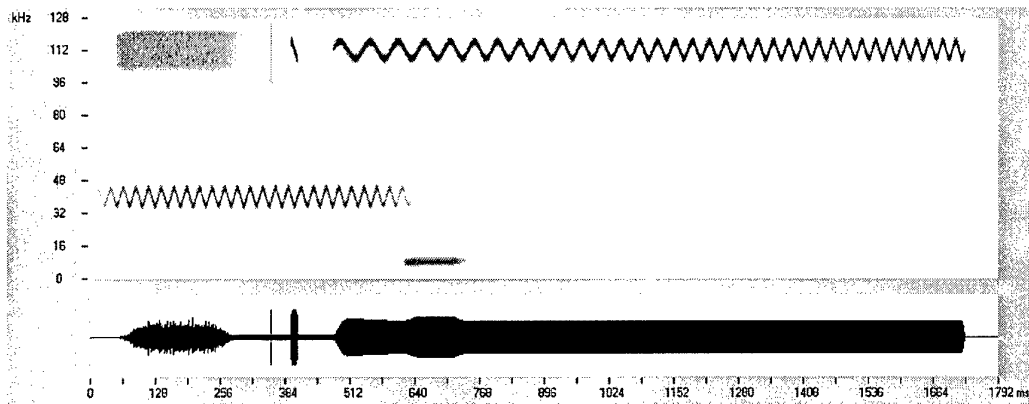


Figure 2. Sequence of test signals created in the range 0-128 kHz (256 ksamples/sec).

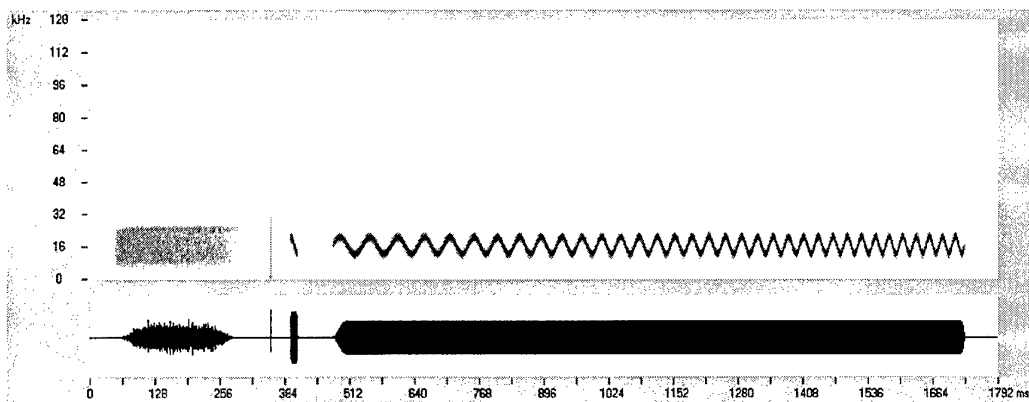


Figure 3. Spectrogram of the test signals after shifting the range 96-128 kHz down to 0-32 kHz.

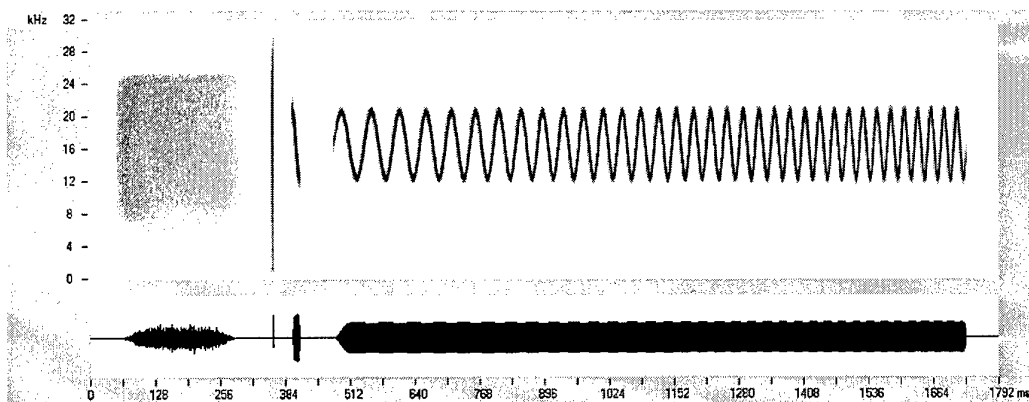


Figure 4. Spectrogram after down shifting (96 kHz) and subsampling (4:1) the test signals.

Platform	Intel Pentium architecture with Windows 95/98/NT/2000/ME
Display	wrap-around or scrolling high resolution real-time spectrogram and cepstrogram
Channels	1 or 2 channels
Sampling rates	up to 192000 samples/sec depending on the installed sound device
Dynamic range	16 bit resolution, 96 dB dynamic range on display
Max Bandwidth	up to 24 kHz with standard sound devices up to 96 kHz with optional PCI cards in desktop PCs
Min Bandwidth	depends on the installed sound device. Digital zooming allows to lower the bandwidth 32 times
Recording size	up to 2GB per file, automatic file splitting to overcome the 2GB limit
Recording format	standard 16 bit uncompressed .wav files (mono or stereo)
Play/Analysis	16 bit uncompressed .wav files and custom files up to 2GB

Hardware requirements:

CPU	at least a P233 for single channel operation. A PII 300 or faster recommended
RAM	64 MB or more
Display	at least 800x600 with 16M colours, 1024x768 or higher recommended
Sound I/O	any standard sound device compatible with Windows, including analog & digital ISA and PCI boards, digital & analog USB and FireWire audio devices, standard notebooks. Multiple sound devices are supported

High Speed Option:

CPU	at least a PII 300, a PIII 500 or faster recommended
Acquisition device	National Instruments DAQ (PCMCIA or FireWire for notebooks, PCI or FireWire for desktops)
Channels	1 channel (up to 8 channels in development)
Sampling rates	500-500k samples/sec with 12 bit resolution (on a PIII 500)

Other Options:

External anti-aliasing filters, radio receiver for sonobuoys, GPS for position logging and time synchronization, 100Mbit ethernet for networked operations, 11Mbit wireless ethernet, wireless audio, external storage units

Table 1 . Summary of features.

4. PRACTICAL APPLICATIONS

High resolution real-time capabilities, typically available in much more expensive instruments, are very useful in behavioural experiments to monitor the acoustic activities of the subjects, for example to find correlations among observed behaviours and emitted/received signals. This allows to immediately evaluate the results of an experiment instead of waiting for later analyses on the recordings; also, this makes easier to analyze long recordings and to access large sound databases. A portable version based on a notebook can be easily moved across laboratories or used on-field.

A typical application for which the software was developed is the monitoring, recording and classification of sounds received by towed arrays or other underwater acoustic sensors while performing marine mammals' acoustic and visual surveys. Lightweight portable equipment is of great value as it allows performing acoustic surveys on small platforms.

Since its earlier DOS versions, the DSPW has been extensively used in our activities in the Mediterranean Sea [6] and in particular in field research when continuous monitoring was required. Typically, we use it to display the signals received by two hydrophones of a towed array while performing passive acoustic surveys for marine mammals on board of sailing boats. In this use, the spectrogram display is also useful to evaluate the quality of the entire electroacoustic chain and to optimize the instrumental setup, array positioning and cruising speed in relation to hydrodynamic and propeller's noise. A number of software and hardware tools have been also developed and tested to improve the management of the system and to make easier the on board activities. Among these, a radio transmitter to broadcast received signals to headphones on the deck, data logging utilities, GPS logging, and a wireless network to share data.

The continuous high resolution display of received sounds, even short and weak signals, can reveal the presence of distant dolphins or sperm whales. By zooming into the low frequency range it is also possible to detect fin whale sounds. Often, acoustic detection largely anticipates visual sightings in daylight surveys, while during night surveys it allows the detection of animals and behaviours otherwise not evident [7]. The comparison between results gained with passive acoustics and visual sightings is still a matter of discussion as each method depends on the equipment and on the observation platform. It is anyway evident that the two methods should be improved and integrated to provide accurate results on censusing while studying marine mammals and their environment.

Recently, the system has been tested as a monitoring tool for the application of acoustic mitigation policies aimed at reducing the impact on marine mammals of underwater high power acoustic sources. Connected to good passive acoustic receivers it can reveal the presence of vocalizing animals within or close to the operation area and thus warn about a possible threat to marine life.

The DSPW was employed during two marine mammal surveys, named SIRENA '99 and SIRENA '00, organized by NATO Saclant Undersea Research Centre within the "Sound, Oceanography and Living Marine Resources" (SOLMAR) project [8,9]. This is a NATO joint research project aimed at increasing the knowledge about marine mammals and their environment to set and test acoustic mitigation policies.

In Sirena '99, other than for continuous monitoring of towed array phones, the DSPW was used to monitor sonobuoys for detecting fin whales sounds by means of real-time zooming in the 0-100 Hz range. During Sirena '00 an unattended recording system was set up to continuously display and record sounds for 8 days, producing 18GB of data a day (2 channels, 48 kHz sampling, 16 bit resolution). Other PCs were used for sound analysis and classification by trained observers 24h/24h to obtain, at the end of the cruise, an accurate summary of all sounds detected [8,9].

5. CONCLUSION

The acquisition of large amounts of data, made possible by widely available digital technologies, generates non-trivial problems when storage, classification and retrieval are concerned. Real-time techniques make easier looking through long recordings, nevertheless it is necessary to develop easy to use and reliable automatic recognition software to automatically browse, analyze and catalogue data acquired during long term monitoring operations. Recognition procedures must be flexible enough to allow the definition of few or many, broad or tight, sound categories and to balance the risk of false detections against the risk of missing events.

The DSPW represent a low cost effective approach to sound analysis and processing in ethological studies as well as in a number of different tasks related to the acoustic monitoring of marine mammals and of the underwater environment in general. Also, its use proved to be of great value for education and training when used with a large sound database.

Further development step will necessarily include the extension to multi-channel analysis and the implementation of automatic recognition capabilities.

6. ACKNOWLEDGEMENTS

We thank Saclant Centre and Angela D'Amico for having made possible the SOLMaR project and Office of Naval Research (ONR) for having granted our research project "Bioacoustic Characterization of the Mediterranean Sea" (ONR Grant N00014-99-1-0709). We also thank the Italian Navy for its support to our research and for the ospitality on board of their ships, helicopters and airplanes.

7. REFERENCES

- [1] PAVAN G., 1992. A portable DSP workstation for real-time analysis of cetacean sounds in the field. *European Research on Cetaceans*, Cambridge (UK), 6: 165-169.
- [2] PAVAN G., 1992. A portable PC-based DSP workstation for bioacoustical research. *Bioacoustics*, 4 (1): 65-66. Abstract.
- [3] PAVAN G., 1995. Low-cost real-time spectrographic analysis of sounds. *Bioacoustics*, 6 (1): 81. Abstract.
- [4] PAVAN G., BORSANI J.F., MANGHI M., PRIANO M., 1996. Interactive Digital Sound Library on cetaceans of the Mediterranean Sea. *European Research on Cetaceans*, 9: 81-84.
- [5] Pavan G., Priano M., Manghi M., Fossati C., 1997. Software tools for real-time IPI measurements on sperm whale sounds. *Proc. Underwater Bio-Sonar and Bioacoustics Symposium. Proc. I.O.A.*, 19 (part 9), Loughborough, UK: 157-164.
- [6] PAVAN G., BORSANI J.F., 1997. Bioacoustic research on cetaceans in the Mediterranean Sea. *Mar. Fresh. Behav. Physiol.*, 30: 99-123.
- [7] Manghi M., Fossati C., Priano M., Pavan G., Borsani J.F., Bergamasco C., 1999. Acoustic and visual methods in the *Odontocetes* survey: a comparison in the Central Mediterranean Sea. *European Research on Cetaceans*, 12: 251-253.
- [8] AA.VV. SIRENA '99. SACLANTCEN CD27
- [9] AA.VV. SIRENA '00. SACLANTCEN CD41